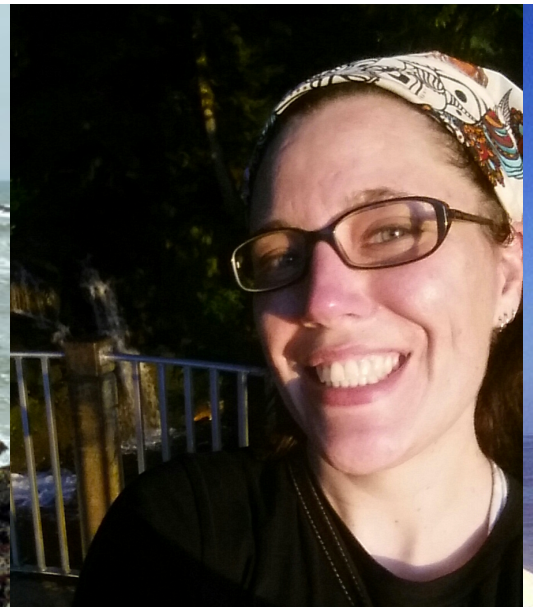
National Center for Genome Analysis Support (NCGAS) use and development of Tripal Genome Browsers on XSEDE's Jetstream

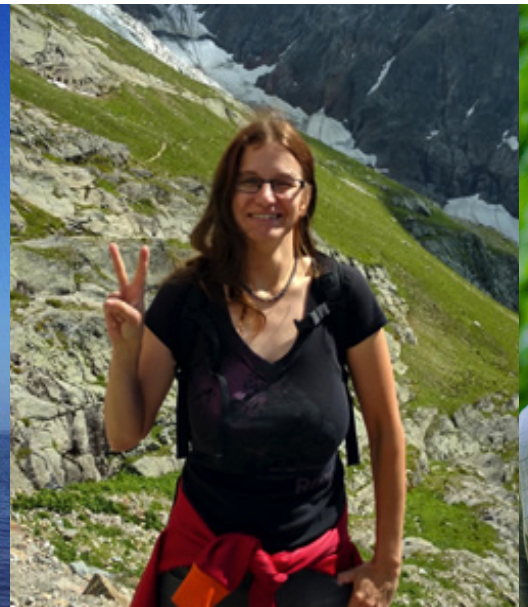
Sheri Sanders
Bioinformatics Analyst
NCGAS @ IU
ss93@iu.edu

What is NCGAS?

The mission of the National Center for Genome Analysis Support is to enable the biological research community of the US to **analyze, understand, and make use of the vast amount of genomic information now available.**

See ncgas.org for all our PAG talks!

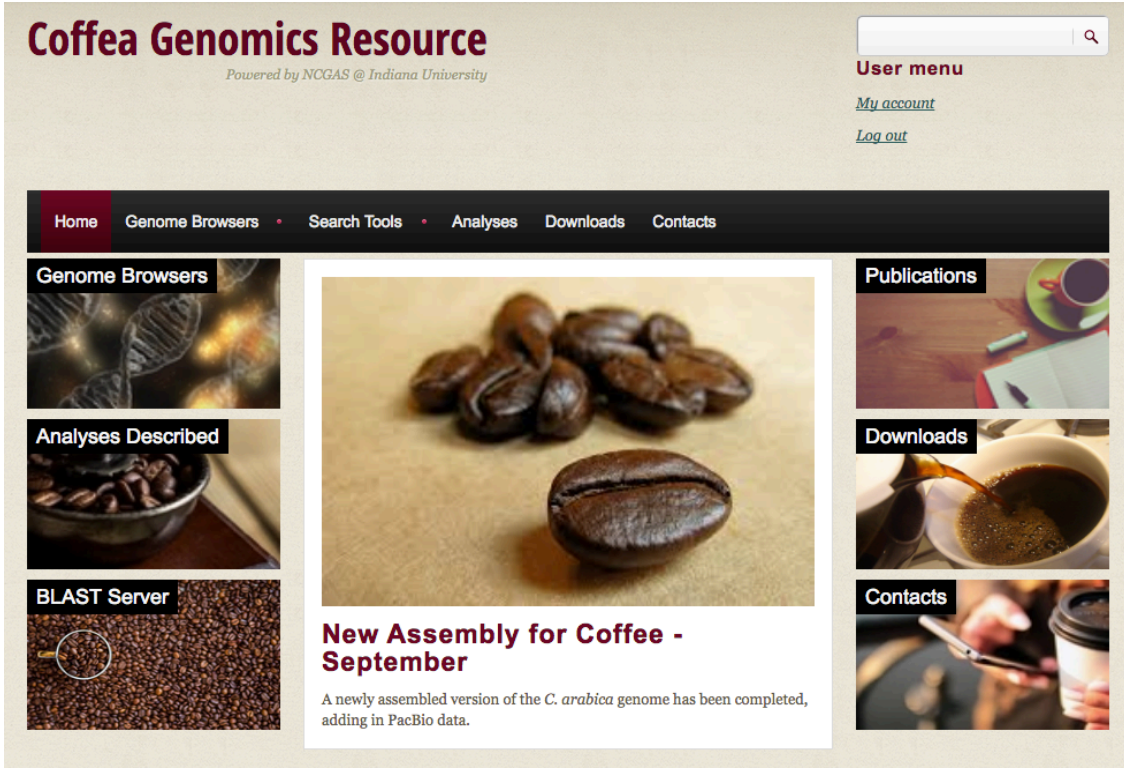



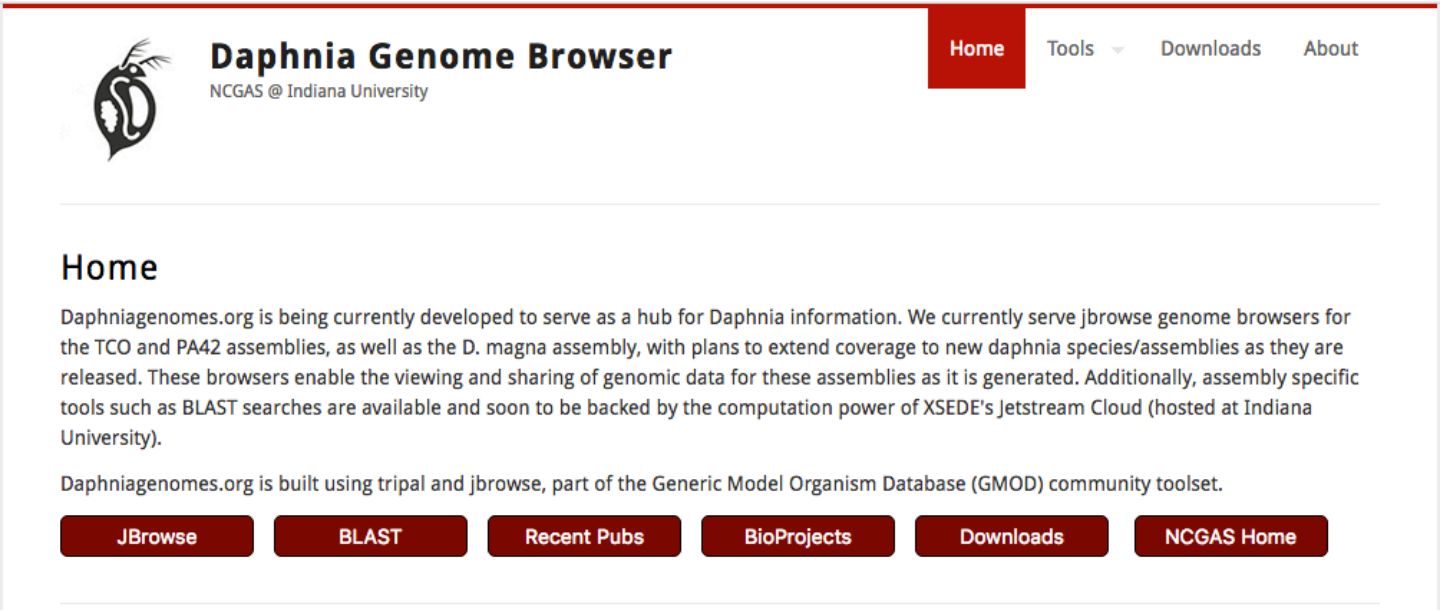


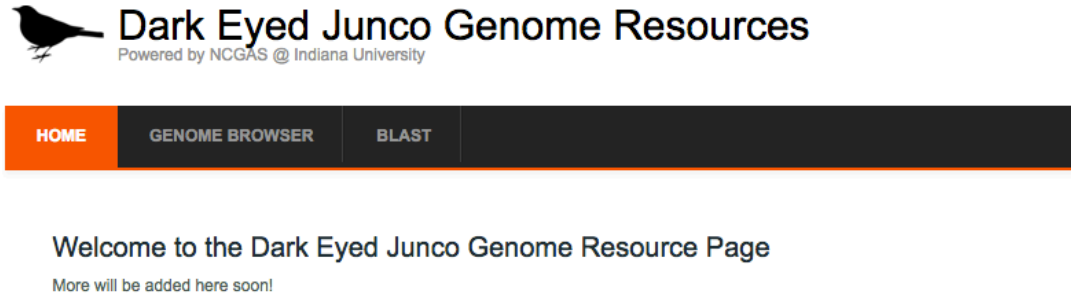

Hosting Browsers

NCGAS has hosted JBrowse instances for years, but in the last ~year, we've moved to Tripal-based genome sites.

Currently, we develop/host resource sites for various projects in daphnia, coffee, peanut, junco, and pumillio frog among others.





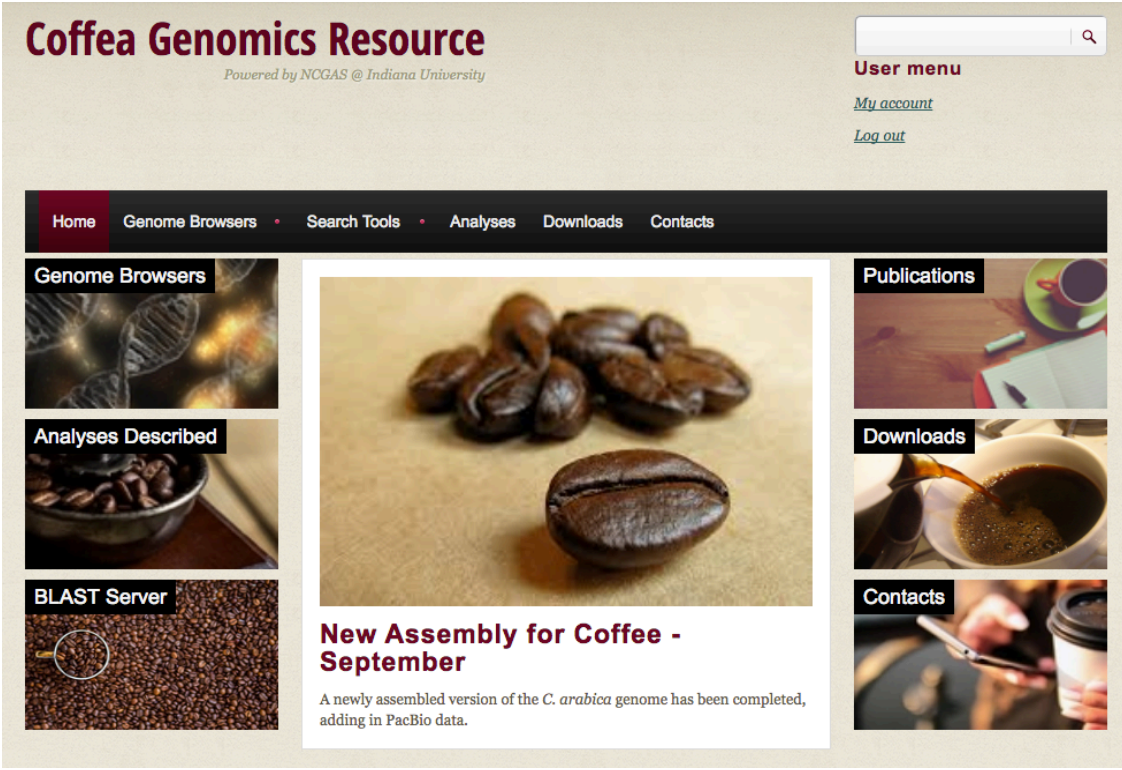


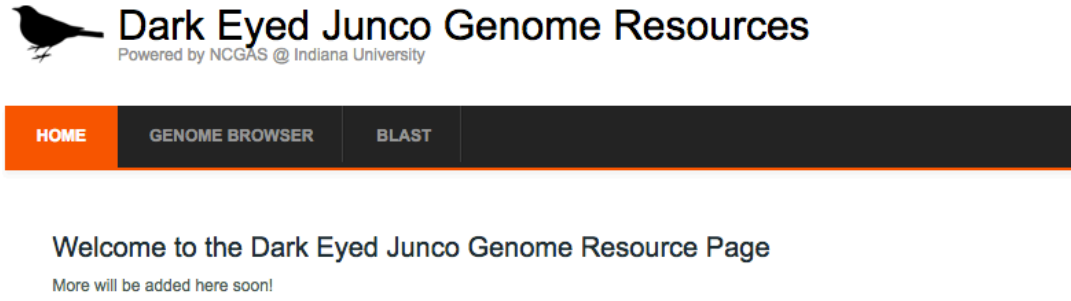
Hosting Browsers

NCGAS has hosted JBrowse instances for years, but in the last ~year, we've moved to Tripal-based genome sites.

Currently, we develop/host resource sites for various projects in daphnia, coffee, peanut, junco, and pumillio frog among others.

Many of groups lack command line skills, so a main goal is to make things automated, more GUI oriented, or just generally easier to implement (in the beginning).





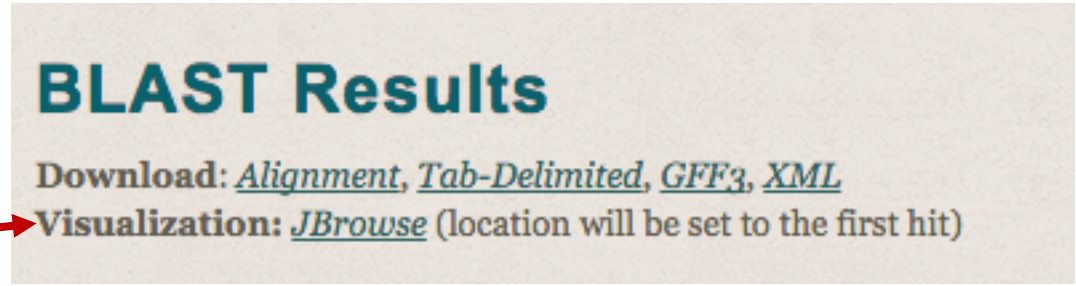
Tools: Tripal_blast tweak

We modified the tripal_blast module to build links to show blast results in internal/external JBrowse sites through URL manipulation.

As a quirk of how it is written, it doesn't require you to own the target JBrowse instance…

For example:

- Blast_UI standard use on home site
- Builds URL based track for hits
- Can be aimed anywhere with the same database…



https://peanutbase.org/node/319881/?data=Aradu1.0&addTracks=[{"label"%3A"blast"%2C"store"%3A"url"%2C"type"%3A"JBrowse%2FView%2FTrack%2FCanvasFeatures"%2C...

PeanutBase

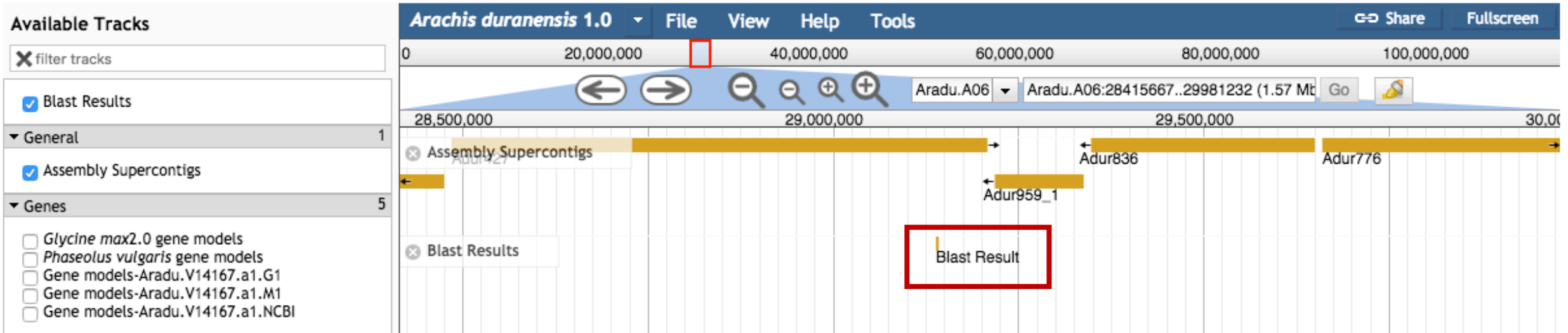
Genetic and genomic data to enable more rapid crop improvement in peanut.

Contact us

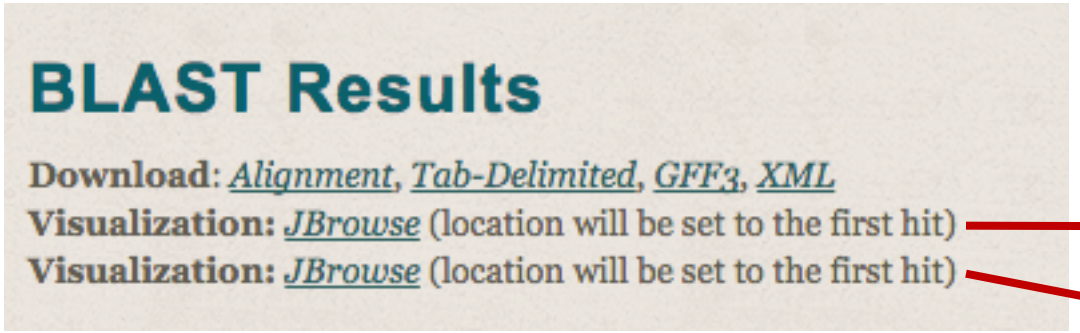
Newsletter signup

Home | Species | Browse & Search | Traits & Maps | Germplasm | Download | Submit Data | Community | External | Help

Show URL for Bookmark



Why would you do this?



To your own browser with unpublished data

To main resource browser with more tracks

Contrived example, but:

- Reduces clutter on your own browser
- Requires less resources to run, without sacrificing data available

Implementation

- Set up is the same as Blast_UI
- Will be available on our GitHub
  - Needs to have configuration added to the GUI
  - Changes are to two files: theme/blast_report.trp.php, and api/blast_ui.api.inc
  - Both updates are marked, current set up requires manual edits
  - README (soon to be help page) walks through how to set it up

DISCLAIMER: This is the first thing I ever modified in Tripal – but it's easy to do, so I use it on all of our sites ^_^.

Tools: Pop Up JB

A GUI-based web tool to spin up new JBrowse instances, with on-the-fly track addition/removal by community, allowing for more flexible community visualization.

Dark Eyed Junco Genome Resources

Powered by NCGAS @ Indiana University

Load a New JBrowse Reference

View | Edit | Grant

If you already have a reference loaded, please proceed to here

Reference Information

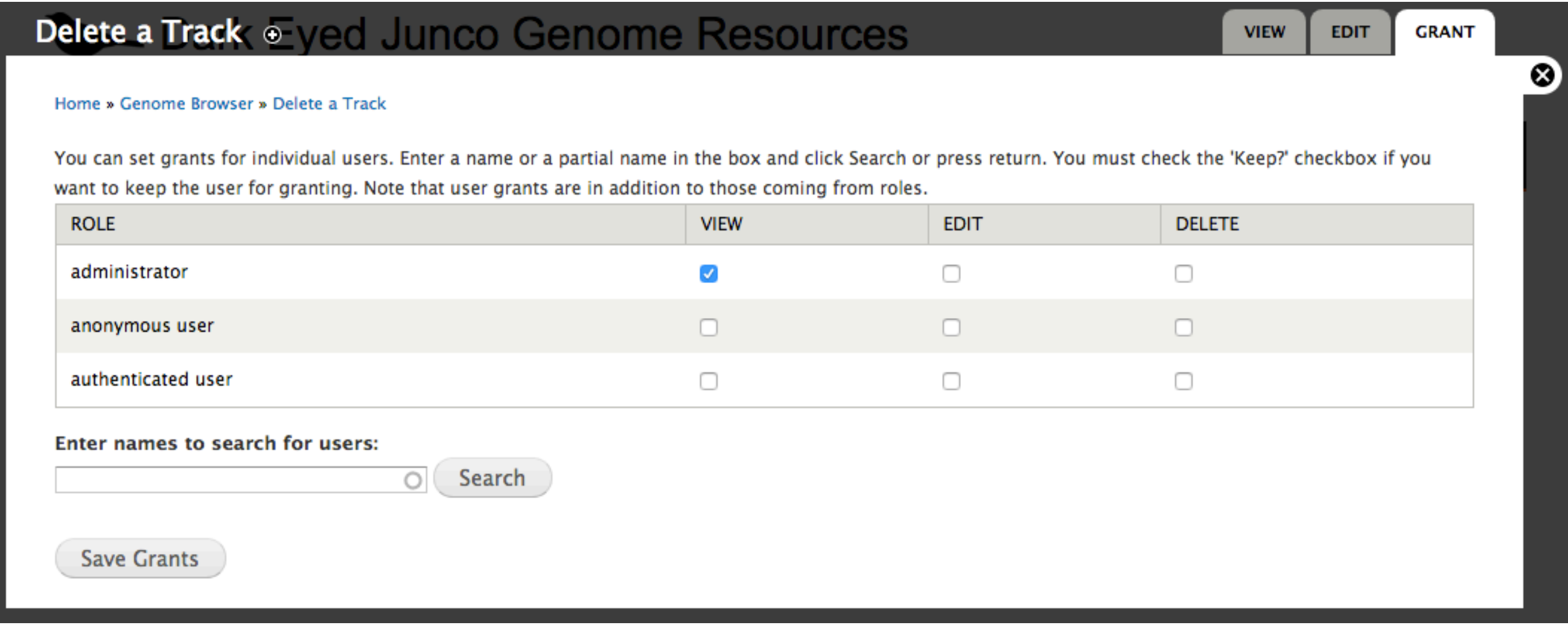
Reference Name (no spaces):

Select file to upload:

Choose File No file chosen

Submit

Security through Node Access



Dark Eyed Junco Genome Resources

Powered by NCGAS @ Indiana University

HOME | GENOME BROWSER | BLAST

Home » Genome Browser » Load a New JBrowse Reference

Load a New JBrowse Reference

View | Edit | Grant

If you already have a reference loaded, please proceed to here

Reference Information

Reference Name (no spaces):

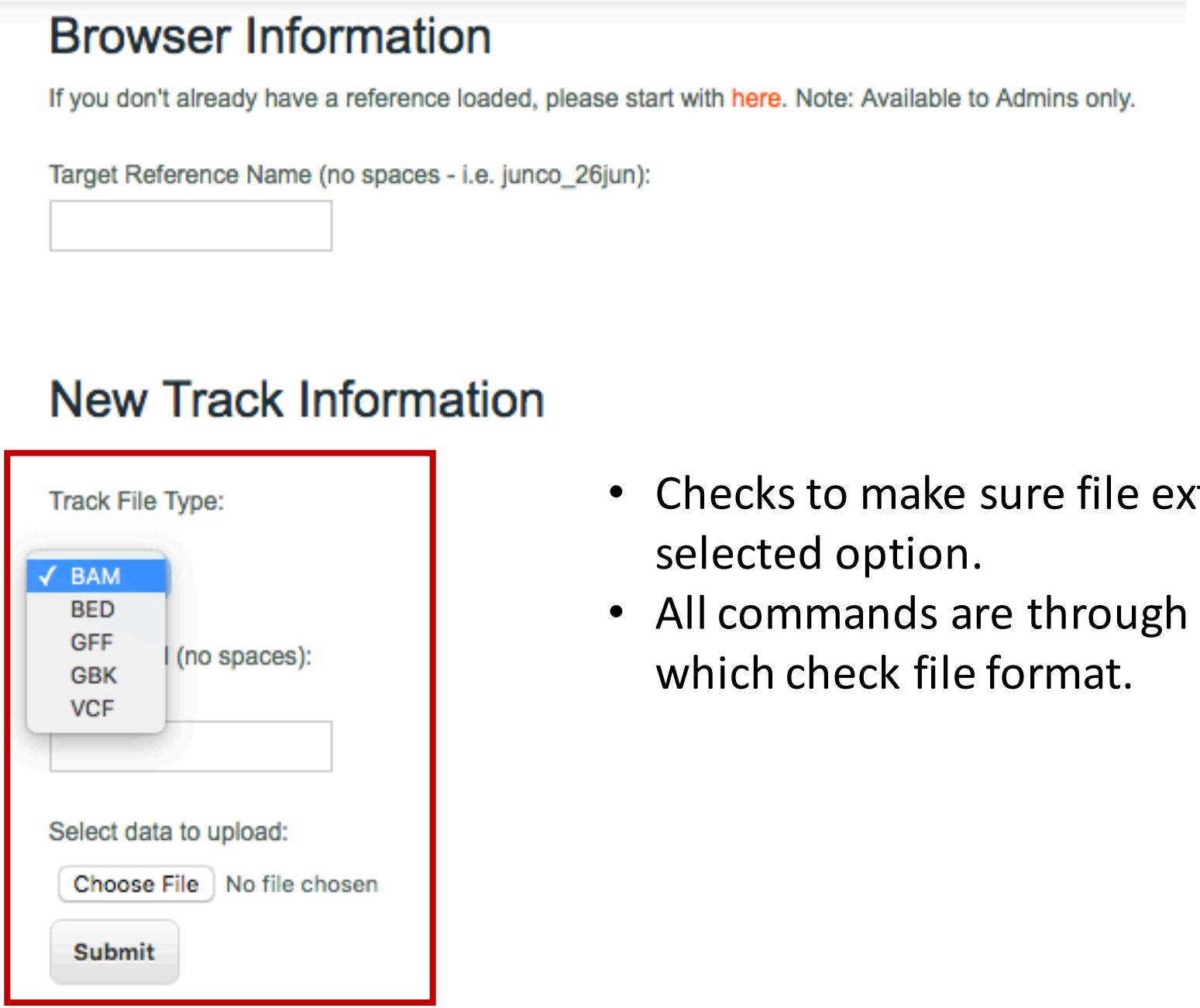
Select file to upload:

Choose File | No file chosen

Submit

Checks if file is *.fa or *.fasta, tells you if/why it fails.

Generally requires increased upload limit in php.ini



- Checks to make sure file extension is appropriate for the selected option.
- All commands are through samtools and JBrowse scripts, which check file format.

Delete a Track

View | Edit | Grant

Browser Information

Reference Name Used (no spaces - listed in browser URL after data =):

Target Track Information

Target Track Label (as listed in browser):

THIS WILL PERMENANTLY DELETE THE TRACK AND DATA. IT IS NOT REVERSABLE!

Submit

Implementation

- Install module as usual
- Have to boost php upload limit
- Have to put an “uploads” folder in JBrowse and open it to apache. I recommend putting a timed purge on uploads/tracks.
- If you are running a frozen version (without track modulation), which is RECOMMENDED, must be non-writable by apache.

- Extensive help page with commands to install dependencies (samtools, bioperl)
- Configure page to define path to JBrowse installation

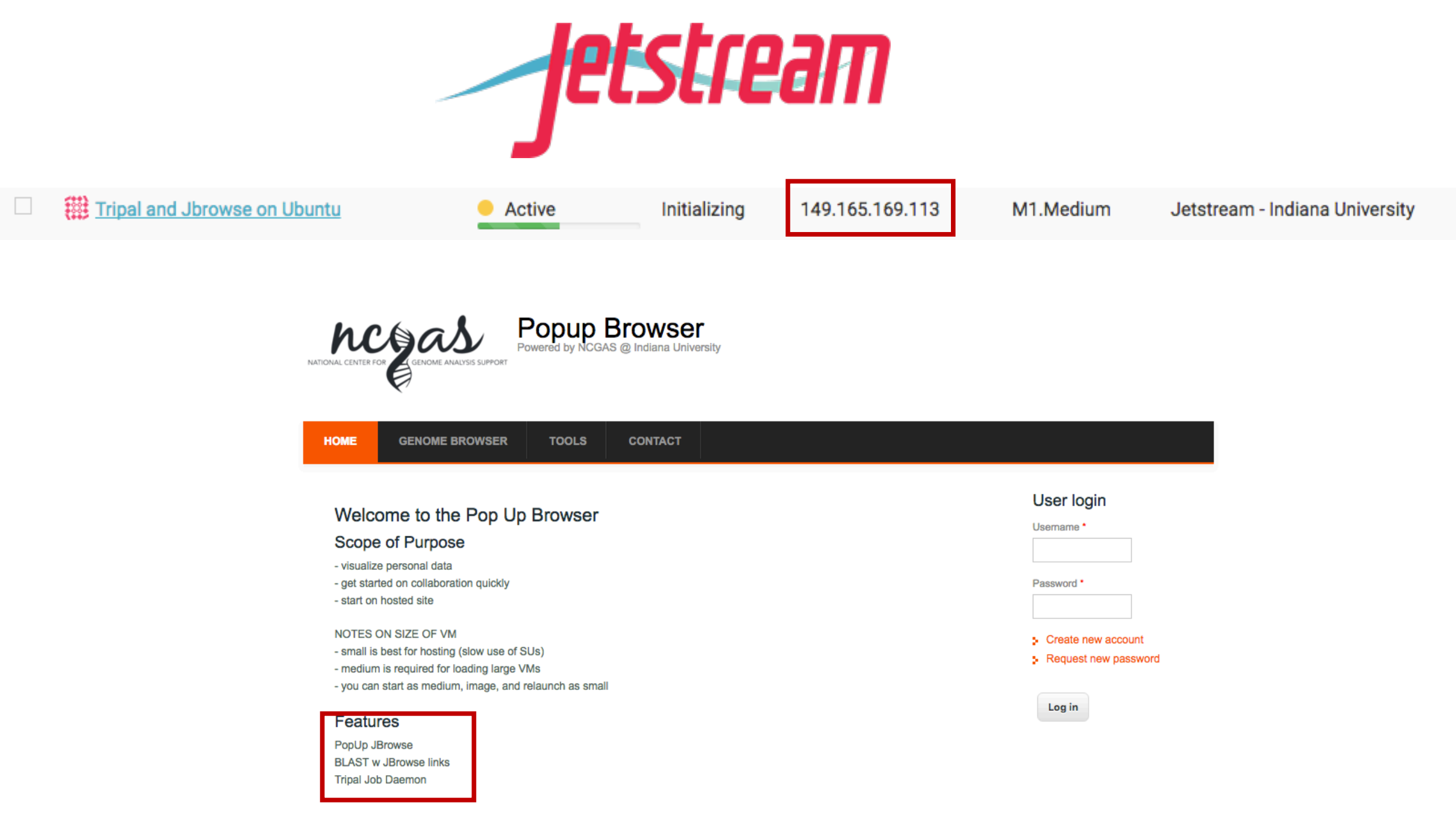
Track flexibility is potentially very useful for popgen

Daphnia project with 96 samples from 96 populations…

- If you want to visualize data for different subsects of that data, this module allows for easy control of authorized users to add and remove tracks to visualize different subsets.
- More attractive than the alternatives of having a highly available admin (not an option in our case) or huge browsers (also not really an option).

Tools: Ready to Run Jetstream VM

We also build a virtual machine image built on the XSEDE cloud (Jetstream) with Tripal, JBrowse, and these tools installed - allowing for free genome browser hosting with minimal command line use and set up.


Jetstream
Tripal and Jbrowse on Ubuntu
Active
Initializing
149.165.169.113
M1.Medium
Jetstream - Indiana University
NATIONAL CENTER FOR GENOME ANALYSIS SUPPORT
Popup Browser
Powered by NCGAS @ Indiana University
HOME
GENOME BROWSER
TOOLS
CONTACT
Welcome to the Pop Up Browser
Scope of Purpose
- visualize personal data
- get started on collaboration quickly
- start on hosted site
NOTES ON SIZE OF VM
- small is best for hosting (slow use of SUs)
- medium is required for loading large VMs
- you can start as medium, image, and relaunch as small
Features
PopUp JBrowse
BLAST w JBrowse links
Tripal Job Daemon
User login
Username *
Password *
Create new account
Request new password
Log in

Set Up and Use

JBrowse:
- see Help for PopUp JB
- once you have built it, add link to GENOME BROWSERS tab (edit the page, fill in the URL and title for the button to work!)

BLAST:
- make a blast reference
- upload it (via globus or jetstream shell/web gui)
- add reference (add content -> Blast Database)

If you are going public:
- change your password!

```
cd /var/www/html/
sudo drush upwd –password="MyNewPassword" admin
```

This is an image – it has a built in password – which needs to be changed if you are going have the site persist.

I do have an account built in for me on the VM, so I can assist with anything anyone needs, as many of our clients are not command line/apache savvy.

This has really helpful in getting people started or running workshop/demos on tripal options without having to walk everyone through installing php/apache/etc.

Also, has made my life much easier, as I can start at a “save point” in the set up process when setting something up for our clients.

Jetstream Access

NCGAS holds three “community” allocations:

- Genomics*
- Field stations*
- Workshops

We help get everyone comfortable with JS and convinced whether or not it will be appropriate for them. Then they can get their own allocations, a process with which we can also help.

Joys of Jetstream

Really good idea to have development and production set ups

- Jetstream SUs get burned faster on larger VMs. Larger VMs are required for more memory – which is needed for **building references/tracks**, but not necessarily displaying them.

- Solution - have a development VM you boot only when needed, use globus to transfer track files between development and production
  - XSEDE users automatically qualify for globus plus (we have tutorials on how to get that)
  - We have walkthroughs on adding globus to a Jetstream instance.

Summary

Tool:	Purpose:
Blast_UI tweak	shows blast results in internal/external JBrowse sites
Popup_JB	on-the-fly track addition/removal by community
Jetstream VM	free genome browser hosting with minimal command line

How to find our tools

- Github (soon)
- Jetstream (by request)
- Email (help@ncgas.org)
- Website will have demos/tutorial to go with JS image soon!